

音声解析に基づいた感情を可視化する手法

櫻井 翔^{*1} 鳴海 拓志^{*2} 佐々木 正人^{*3}

Emotion visualization method based by voice analysis

Sho Sakurai^{*1}, Takuji Narumi^{*2} and Masato Sasaki^{*3}

Abstract

現在の記録メディアは、感情の記録をおこなう際は言語に依存する場合が多く、記録者が感じたことを意図通りに第三者が受け取るとは限らない。しかし、感情を生体情報に関連したパラメータとして記録し、可視化して提示することができれば、出来事そのものだけでなく出来事の中で抱いた感情を追体験することができ、記録者の過去の体験における感情を第三者に向けて伝達、共有がより容易になると考えられる。本稿では、感情の追体験を目的として、音声情報を用いて感情を解析して記録し、そのデータを用いて感情を可視化するビューアについて報告する。

In present, when people records their feelings, record media is often depends on the language, and the third party doesn't receive what the recorder felt according to the intention. However, it is thought that it becomes easier to transmit and to share by being able to relive not only the event but also feelings held in the event, and turning feelings in a past experience of the recorder to the third party if it is possible to present it by recording feelings as a parameter related to living body information, and making it to visible. In this text, we reports on the viewer to make feelings visible by analyzing feelings by using tonal information recording and using the data to relive feelings.

Keywords: emotion, visualization, relive, lifelog and Sound analysis

1. はじめに

記録という行為の背景には、記憶の保存、再生を通じて、自己あるいは他者の体験の追体験をおこない、経験や知識、感動の共有をおこないたいという欲求がある^[1]。

現在の主流な記録メディアは、出来事や経験の内容そのものを記録することは容易であるのに対し、記録者の感情や思考といった主観的な部分の記録をリアルタイムで保存し、再生をおこなうには向いていない。映像や写真は客観的な視点でしか記録できず、個人の主観は状況や人の表情などから推測するしかない。また、日記などは一度体験を忘れてしまうと、忘れてしまった部分は記録として残すことができない。

一方で、映像や写真に感情データを付与して体験を記録するためのインタフェースの研究^[2,3]もおこなわれている。これらの研究では、記録時に同時に感想を述べたり、映像にいくつかの感情を表現する語

彙を選択してタグ付けし、インデキシングするなど感情表現には主に言語が使用されている。言語は自分の主観を他者に伝えるために重要な手段である反面、語彙の受け取り方は人それぞれであり、記録者の意図する通りに感情状態を他者に伝達し、共有できるような記録をおこなうには、記録者による詳細な説明が必要であるという面もある。

これに対し、感情をあるパラメータと捉えて記録し、視覚化して再生することができれば、記録者の主観の第三者へ伝達、共有がより容易にできると考えられる。

言語や表面上の見た目に依存せずに感情を記録し可視化するために、感情の変化にともなって起こる生体反応を利用することが有効であると考えられる。感情の変化とともに変化する生体反応としては表情筋肉の動きや脳波の変化などが挙げられる^[4]が、われわれは出来事の中で発せられる音声情報に注目した。人と会って話すことは、日常生活において高い確率で起こるイベントであるが、個人が単独で存在するときよりも他者の存在があることで感情の表出が促進される^[1]。特に、ともに存在する他者が友人など仲の良い相手である場合、コミュニケーション欲求が高まることにより、感情の表出が促進されるという研究結果が先行研究により示されている。個人で行動するときよりも対人イベントが起こるとき

*1: 東京大学大学院 学際情報学府

*2: 東京大学大学院 工学系研究科/日本学術振興会特別研究員

*3: 東京大学大学院 教育学研究科

*1: Graduate School of Interdisciplinary Information Studies, The University of Tokyo

*2: Graduate School of Engineering, The University of Tokyo/JSPS

*3: Graduate School of Education, The University of Tokyo

の方が感情の変化する可能性が高いと考え、音声情報を利用することを考えた。

本論文では、音声を周波数解析することにより感情状態を数値化し、その変化を分かりやすく可視化するための感情記録ビューアを作成した。そして実際に感情状態が可視化されているか、実験をおこなった。

2. 音声解析による感情の判別

2.1 システムの概要

本稿で述べる感情記録ビューアの概要を図2に示す。ビューアはPCとマイク、ウェブカメラで構成されており、マイクから取り込んだ音をPCで解析するとともに、ウェブカメラで写真を撮影し、時間軸上に同時に並べていく。

2.2 実装

2.2.1 音声解析による感情変化のパラメータと感情値の設定

本システムは、感情をポジティブとネガティブに2極化し、その起伏を表示することで可視化をおこなう。

マイクを通して取り込んだ音声をFFT解析して、音声の強さ、高さ、強さの時間差分、高さの時間差分の4つの指標を設定した。また、感情の高まりの度合いを表すmood値を設定した。mood値は50をニュートラルな状態とし、最高値を100、最低値を0としている。

なお、パラメータやmood値は、何度か事前に実験をおこなった上で、実際の感情変化に沿うと考えられるよう、以下のような設定をおこなった。

得られた音声をFFT解析し、0~860Hzを43Hzごとに20個の帯域に分けた。各帯域の音の振幅の平均値を a_1, a_2, \dots, a_{20} とし、全体の振幅の平均値を音の強さ A_n として、(1)の式で求める。

$$A_n = \sum_{n=1}^{20} a_n \quad (1)$$

また、各帯域の音の周波数の平均値を f_1, f_2, \dots, f_{20} とし、全体の周波数の平均値を音の高さ F_n として、(2)の式で求める。

$$F_n = \sum_{n=1}^{20} \frac{f_n a_n}{A_n} \quad (2)$$

さらに、1/60秒ごとに得られた音の強さを順に A_1, A_2, \dots, A_n として、音の強さの時間差分の絶対値 Ad を(3)の式で求める。

$$Ad_n = \{A_n - A_{n-1}\}^2 \quad (3)$$

同様に、音の高さの時間差分の絶対値は(4)の式で求める。

$$Fd_n = \{F_n - F_{n-1}\}^2 \quad (4)$$

以上の4つの数値から、 $A_n > 0.8$ であれば音声が入力されている状態とし、感情の変化をあらゆる値moodを M_n とおく。そして、 $A_n < N_1$, $Ad_n \leq N_2$ のときはネガティブな状態と判断し、さらに $f_n < P_1$, $Fd_n \leq P_2$ であればゆるやかに、 $f_n \geq P_1$, $Fd_n > P_1$ であれば急激にネガティブな状態へ変化すると判断して、 M_n はそれぞれ(5)、(6)の様に求めた。

$$M_n = M_{n-1} - \frac{a_n}{10} \quad (5)$$

$$M_n = M_{n-1} - \frac{a_n}{15} \quad (6)$$

逆に、 $A_n > N_1$, $Ad_n > N_2$ のときはポジティブな状態と判断し、 M_n は(7)の式で求める。

$$M_n = M_{n-1} + \frac{a_n}{5} \quad (7)$$

このとき、事前におこなった実験により、 $N_1=75$, $N_2=12$, $P_1=20$, $P_2=8$ とした、また、ネガティブな状態と判断したときのように場合分けをしなかった理由は、ポジティブな状態と判断されたときは全体的に急激に変化し、ゆるやかと取られる状態が見られなかったためである。

これらは、高めの声、大きな声、抑揚がはっきりした話し方の場合はmood値が50から上昇してポジティブと判断される。反対に、低めの声、小さい声、抑揚があまりない話し方の場合は50から下降してネガティブであるという判断をおこなう。

2.2.2 Weather-vox central

ここで、本システム作成の前に試作した作品について少し触れておきたい。筆者らは、2008年7月に東京大学で開催された制作展 iii Exhibition 9にて、『weather-vox central』という作品を出展した(図1, 2)。この作品は、音声の周波数を解析して、天気の変化で感情状態を表現する作品である。感情がポジティブと判断された場合は空が晴れ、ネガティブと判断された場合は夜になっていく。ビューア作成の際に使用しているパラメータや値は、『weather-vox central』制作の際に使用したものと同一である。

実際に来場者に作品を体験してもらったときに、「明日はディズニーランドに行く」という発言の際にはmood値が一気に画面が晴れたり、「課題が終わらない」という発言の際にはゆっくりと夜になっていくなど、ある程度予想された反応が見られた。一方で、次の日の仕事が憂鬱だという来場者が「明日の仕事が楽しみだ」とわざと明るい声で発言した際、晴天が途端に夜になるといった場面も見られた。他

3. 評価実験

今回、本システムを作成して何度か実験をおこなったが、ここでは 2008 年 10 月 2 日におこなった実験について述べる。

Fig. 6 は、ある展示のためのミーティングの際に記録したビューアの画像である。このミーティングには 7 人が参加し、それぞれが各自の作品について発表をおこなった。各人が発表をする際に本記録システムを起動し、感情の変化の記録をおこない、のちにその変化を可視化し、記録者に提示することで実際の感情変化との比較をおこなった。

作品についての説明のときは話し方が単調であり、mood 値の変化はほとんど見られなかったのに比べ、展示会場の現地の説明や、特に力を入れて説明している部分では mood 値は上昇した。特に笑いが起きたときの mood 値の上昇は顕著であった。逆に説明に詰まるなど、焦りが見えたときや注意喚起をおこなったときは mood 値は下降し、決して気分が良い状態ではないと判断されていることが分かる。このような変化は、どの被験者が発表するときもおおむね同様の結果が得られた。

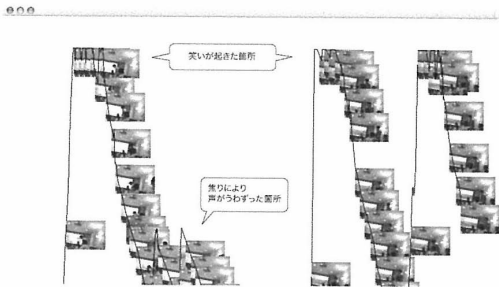


図 6 2008 年 10 月 2 日の記録の一部
Fig.6 A part of record October 6, 2008

4. まとめ

本稿では、感情の追体験を目的として、感情を可視化して記録するビューアを提案した。記録システムでは、カメラからの画像とともに、マイクに入力される音声を周波数解析し、その結果によってポジティブ/ネガティブの 2 極間での感情状態の変化を記録する。ビューアでは、記録された感情をあらゆるパラメータをグラフであらわし、同時に記録された写真を提示することで、客観的事実と感情を同時に提示できるような実装をおこなった。

このシステムを用いて、日常生活の中で感情状態の変化を記録し可視化できるか実際に実験をおこなったところ、複数の被験者において感情状態の変化

を可視化できていることを確認することができた。

しかし、感情の追体験という側面においては十分な実験をおこなうことができず、本稿で提案したビューアによる記録が実際に追体験に有効であるかは、厳密に条件を設定した実験と細かい考察が必要である。また、ポジティブとネガティブという 2 極値だけでなく、たとえば喜怒哀楽の様な基本感情の差異を表現する必要があるだろう。

感情の追体験を目的として、今後は以上に挙げた課題を含め、感情状態をより細かく判定し、可視化するための方法を考えていきたい。また、感情状態の判定に今回は音声情報を利用したが、他に有効と考えられる生体情報などを探してデータとして利用することも考えられる。

謝辞

本稿で紹介したシステムの作成には、鈴木隆志、栗山貴嗣の両氏に、プログラムの構築に際し多大な協力を得た。この場を借りて、両氏に心より感謝します。

参考文献

- [1] 北村英哉, 木村晴: 感情研究の新展開; ナカニシヤ出版 (2006)
- [2] 角康之, 河村竜幸: 体験メディアの構築に向けて: 体験の記録・利用の技術動向; 人工知能学会第 20 回全国大会 (2006)
- [3] 志村将吾, 平野靖, 梶田将司, 間瀬健二: 体験映像への感情付与インタフェース; 情報処理学会, 2005-HI-115-(11), Vol.2005, No.95, pp61-68 (2005)
- [4] 高橋雅延, 谷口高士: 感情と心理学-発達・生理・認知・社会・臨床の接点と新展開; 北大路書房 (2002)